Technische Universität Chemnitz

Sonderforschungsbereich 393

Numerische Simulation auf massiv parallelen Rechnern

Peter Benner   Ralph Byers   Volker Mehrmann   Hongguo Xu

A Unified Deflating Subspace Approach for Classes of Polynomial and Rational Matrix Equations

Preprint SFB393/00-05

Preprint-Reihe des Chemnitzer SFB 393

SFB393/00-05 **January 2000**

Contents

Authors:

Peter Benner
Zentrum für Technomathematik,
Fachbereich 3/Mathematik und Informatik
Universität Bremen
D-28334 Bremen, Germany
benner@math.uni-bremen.de

Ralph Byers
Department of Mathematics
University of Kansas
Lawrence, KS 66045
USA
byers@math.ukans.edu

Volker Mehrmann
Fakultät für Mathematik
Technische Universität Chemnitz
D-09107 Chemnitz
Germany
mehrmann@mathematik.tu-chemnitz.de

Hongguo Xu
Deparment of Mathematics
Case Western Reserve University
10900 Euclid Avenue
Cleveland, Ohio, 44106, USA
hxx7@po.cwru.edu

A Unified Deflating Subspace Approach for Classes of Polynomial and Rational Matrix Equations[1]

Peter Benner Ralph Byers[2] Volker Mehrmann Hongguo Xu

Abstract

A unified deflating subspace approach is presented for the solution of a large class of matrix equations, including Lyapunov, Sylvester, Riccati and also some higher order polynomial matrix equations including matrix m-th roots and matrix sector functions. A numerical method for the computation of the desired deflating subspace is presented that is based on adapted versions of the periodic QZ algorithm.

Keywords. Eigenvalue problem, deflating subspace, Lyapunov equation, Sylvester equation, Riccati equation, matrix roots, matrix sector function, periodic QZ algorithm.
AMS subject classification. 65F15, 93B40, 93B36, 93C60.

1 Introduction

The relationship between matrix eigenvalue problems and the solution of polynomial or rational matrix equations has been an important research topic in numerical linear algebra due to its many applications, for example in control theory, see e.g., [3, 16, 21, 24, 27, 28, 38, 39, 45, 55]. It is well known that many polynomial or rational matrix equations can be solved by computing invariant subspaces of matrices and deflating subspaces of matrix pencils. Examples include Schur methods for matrix m-th roots, sector functions, algebraic Riccati equations, Sylvester equations, Lyapunov equations and their generalizations [4, 6, 14, 17, 19, 20, 22, 23, 25, 30, 35, 40, 41, 45, 47, 51].

In this paper we consider the computation of deflating subspaces of a (generalized) matrix pencil of the form $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ with complex $n \times n$ matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. (There exist similar methods for real matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ that use only real arithmetic. However, for ease of presentation, we will present the complex case only.) We show that from these deflating subspaces the solution of many classes of matrix equations can be obtained. These include linear and quadratic matrix equations as well as some rational and higher order polynomial matrix equations like the matrix m-th root and m-th sector function, see [5, 33, 35, 49].

[1]All authors were partially supported by *Deutsche Forschungsgemeinschaft*, Sonderforschungsbereich 393, “Numerische Simulation auf massiv parallelen Rechnern”.

[2]This author was partially supported by National Science Foundation awards CCR-9732671, MRI-9977352, and by the NSF EPSCoR/K*STAR program through the Center for Advanced Scientific Computing.

2 Preliminaries

By $\mathbf{C}^{n\times n}$ and $\mathbf{R}^{n\times n}$ we denote the sets of complex or real $n\times n$ matrices, respectively. By I_n and 0_n the $n\times n$ identity matrix and zero matrix, respectively and we set $J_n = \left[\begin{smallmatrix} 0_n & I_n \\ -I_n & 0_n \end{smallmatrix}\right]$. We omit the subscript n, if the sizes are clear from the context.

In the paper we will consider eigenvalue problems, i.e., the computation of eigenvalues and deflating subspaces for matrix pencils of the form $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ with $n\times n$ matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$.

Definition 2.1 Consider the matrix pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ with $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbf{C}^{n\times n}$.

1. If $\det(\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C})$ is not identically zero, then the matrix pencil is said to be *regular*.
2. The generalized eigenvalues of the pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ are the pairs $(\alpha, \beta) \in \mathbf{C}^2 \setminus \{(0, 0)\}$ for which $\det(\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}) = 0$. If (α, β) is an eigenvalue with $\alpha \neq 0$, then it is said to be finite eigenvalue and it is often identified with the number $\lambda = \beta/\alpha$. If $(0, \beta)$ is an eigenvalue, then it is said to be an infinite eigenvalue.
3. A k–dimensional subspace $\mathbf{U}$ is called *right deflating subspace* of the regular matrix pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ if for a full rank matrix $U \in \mathbf{C}^{n\times k}$ with $\operatorname{range} U = \mathbf{U}$, there exist a full rank matrix $V \in \mathbf{C}^{n\times k}$ and $R_A, R_{BC} \in \mathbf{C}^{k\times k}$ such that $\mathcal{A}U = VR_A$ and $\mathcal{B}\mathcal{C}U = VR_{BC}$. (Regularity of $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ implies the regularity of $\alpha R_A - \beta R_B R_C$.)
4. A k–dimensional subspace $\mathbf{U}$ is called *left deflating subspace* if it is a right deflating subspace of $\alpha\mathcal{A}^H - \beta\mathcal{C}^H\mathcal{B}^H$.
5. A k–dimensional subspace $\mathbf{W}$ is called an *interior deflating subspace* of the regular matrix pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ if for a full rank matrix $W \in \mathbf{C}^{n\times k}$ with $\operatorname{range} W = \mathbf{W}$, there exist matrices $U, V \in \mathbf{C}^{n\times k}$ and R_A, $R_B, R_C \in \mathbf{C}^{k\times k}$ such that $\mathcal{A}U = VR_A$, $\mathcal{B}W = VR_B$ and $\mathcal{C}U = WR_C$.

Note that if $\mathcal{C} = I$, then an interior deflating subspace is just a classical right deflating subspace of $\alpha\mathcal{A} - \beta\mathcal{B}$ and if $\mathcal{B} = \mathcal{C} = I$, then the subspaces are usually called *right and left invariant subspaces* of the matrix $\mathcal{A}$.

We denote by $\Lambda(\mathcal{A})$ the spectrum of a square matrix $\mathcal{A}$ and analogously by $\Lambda(\mathcal{A}, \mathcal{B}\mathcal{C})$ the set of generalized eigenvalues of the pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$. For such pencils a generalized periodic Schur form and a periodic QZ algorithm to compute it were introduced in [15, 29].

Proposition 2.2 *For a matrix pencil $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ with $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbf{C}^{n\times n}$, there exist unitary matrices $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$, such that the matrices $\mathcal{V}^H\mathcal{A}\mathcal{U}$, $\mathcal{V}^H\mathcal{B}\mathcal{W}$ and $\mathcal{W}^H\mathcal{C}\mathcal{U}$ are all upper triangular. The generalized eigenvalues of the pencil are displayed by the diagonal entries of the three triangular matrices and can be obtained in any desired order by an appropriate choice of $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$.*

(If $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are real matrices, then there exists a similar generalized periodic Schur form involving quasi-triangular matrices.)

The relationship between deflating subspaces and large classes of matrix equations is described in the following proposition.

Proposition 2.3 *Consider a matrix pencil* $\alpha\mathcal{A} - \beta\mathcal{B}\mathcal{C}$ *with matrices* $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbf{C}^{n\times n}$. *Partition the matrices* $\mathcal{A}$, $\mathcal{B}$ *and* $\mathcal{C}$ *in* m *compatible blocks* $\mathcal{A} = [A_{i,j}]$, $\mathcal{B} = [B_{i,j}]$ *and* $\mathcal{C} = [C_{i,j}]$ *with blocks* $A_{ii}, B_{ii}, C_{ii} \in \mathbf{C}^{n_i\times n_i}$, $i = 1, \ldots, m$. *Suppose that there exist matrices*

$$U = \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix}, \quad V = \begin{bmatrix} V_1 \\ \vdots \\ V_m \end{bmatrix}, \quad W = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix},$$

with $n_i \times n_1$ *blocks* U_i, V_i *and* W_i *along with* $n_1 \times n_1$ *matrices* R_A, R_B *and* R_C *such that*

$$\mathcal{A}U = VR_A, \quad \mathcal{B}W = VR_B, \quad \mathcal{C}U = WR_C. \tag{1}$$

(i) If U_1, V_1 *and* W_1 *are nonsingular, then the matrices* $X_i := U_{i+1}U_1^{-1}$, $Y_i := V_{i+1}V_1^{-1}$, $Z_i := W_{i+1}W_1^{-1}$, $i = 1, \ldots, m-1$ *satisfy the matrix equations*

$$A_{k,1} + \sum_{i=1}^{m-1} A_{k,i+1}X_i = Y_{k-1}(A_{1,1} + \sum_{i=1}^{m-1} A_{1,i+1}X_i), \tag{2}$$

$$B_{k,1} + \sum_{i=1}^{m-1} B_{k,i+1}Z_i = Y_{k-1}(B_{1,1} + \sum_{i=1}^{m-1} B_{1,i+1}Z_i), \tag{3}$$

$$C_{k,1} + \sum_{i=1}^{m-1} C_{k,i+1}X_i = Z_{k-1}(C_{1,1} + \sum_{i=1}^{m-1} C_{1,i+1}X_i), \tag{4}$$

for $k = 2, \ldots, m$.

(ii) If the matrices $\{X_i\}_{i=1}^{m-1}$, $\{Y_i\}_{i=1}^{m-1}$ *and* $\{Z_i\}_{i=1}^{m-1}$ *satisfy the matrix equations (2)–(4), then the matrices*

$$U = \begin{bmatrix} I \\ X_1 \\ \vdots \\ X_{m-1} \end{bmatrix}, \quad V = \begin{bmatrix} I \\ Y_1 \\ \vdots \\ Y_{m-1} \end{bmatrix}, \quad W = \begin{bmatrix} I \\ Z_1 \\ \vdots \\ Z_{m-1} \end{bmatrix}, \tag{5}$$

satisfy (1) with

$$R_A = A_{1,1} + \sum_{i=1}^{m-1} A_{1,i+1}X_i, \quad R_B = B_{1,1} + \sum_{i=1}^{m-1} B_{1,i+1}Z_i, \quad R_C = C_{1,1} + \sum_{i=1}^{m-1} C_{1,i+1}X_i.$$

Proof. The proof of the first part follows by elementary calculations, comparing the corresponding blocks on both sides of (1) and using the nonsingularity of the matrices U_1, V_1 and W_1. The second part is immediate. □

Remark 2.4 The equations in (2)–(4) are matrix equations in the matrix variables $\{X_i\}_{i=1}^{m-1}$, $\{Y_i\}_{i=1}^{m-1}$ and $\{Z_i\}_{i=1}^{m-1}$. Specific cases that we study below are given by choosing m and appropriate blocks $A_{i,j}$, $B_{i,j}$ and $C_{i,j}$.

Remark 2.5 Equations (2)–(4) may have many solutions. But for each set of solutions, the matrices U, V, W as in (5) determine a deflating subspace associated with the matrix subpencil $\alpha R_A - \beta R_B R_C$. As we see from Proposition 2.3, for the converse we need the nonsingularity of U_1, V_1 and W_1. This implies that for a matrix pencil the deflating subspaces may exist (they always exist when the pencil is regular), but the solution of the related matrix equation may not exist.

In the following sections we study in more detail special cases of matrix equations as in Proposition 2.3.

3 Quadratic matrix equations

We first study quadratic matrix equations which arise from the case $m = 2$ in Proposition 2.3. In this case

$$\mathcal{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}, \tag{6}$$

and

$$U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}, \quad W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}. \tag{7}$$

The matrix equations (2)–(4) then take the form

$$\begin{aligned} A_{21} + A_{22}X &= Y(A_{11} + A_{12}X), &(8)\\ B_{21} + B_{22}Z &= Y(B_{11} + B_{12}Z), &(9)\\ C_{21} + C_{22}X &= Z(C_{11} + C_{12}X). &(10) \end{aligned}$$

These equations can be viewed as generalized Lur'e equations [32].

As a corollary of Proposition 2.3 we have the following result.

Corollary 3.1 *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be as in (6). Let U, V and W be as in (7) and assume that they satisfy*

$$\mathcal{A}U = VR_A, \quad \mathcal{B}W = VR_B, \quad \mathcal{C}U = WR_C \tag{11}$$

for some square matrices R_A, R_B and R_C. If U_1, V_1 and W_1 are invertible, then $X = U_2U_1^{-1}$, $Y = V_2V_1^{-1}$ and $Z = W_2W_1^{-1}$ satisfy (8)–(10). Conversely, if X, Y and Z satisfy (8)–(10), then $U = \left[\begin{smallmatrix} I \\ X \end{smallmatrix}\right]$, $V = \left[\begin{smallmatrix} I \\ Y \end{smallmatrix}\right]$ and $W = \left[\begin{smallmatrix} I \\ Z \end{smallmatrix}\right]$ satisfy (11) with

$$R_A = A_{11} + A_{12}X, \quad R_B = B_{11} + B_{12}Y, \quad R_C = C_{11} + C_{12}Z.$$

Multiply (9) from the right by $C_{11} + C_{12}X$, rearrange the equation and use (10) to obtain

$$(B_{22} - YB_{12})(C_{21} + C_{22}X) = (YB_{11} - B_{21})(C_{11} + C_{12}X). \tag{12}$$

If $\mathcal{D} = \mathcal{B}\mathcal{C} = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, then the system takes the form

$$
\begin{aligned}
YA_{12}X - A_{22}X + YA_{11} - A_{21} &= 0, && (13)\\
YD_{12}X - D_{22}X + YD_{11} - D_{21} &= 0. && (14)
\end{aligned}
$$

For the solution of (13)–(14) we do not need the nonsingularity of W_1. For completeness we state this special case as a corollary.

Corollary 3.2 *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be as in (6) and let $\mathcal{D} = \mathcal{B}\mathcal{C}$. Let U and V be as in (7) and satisfy*

$$\mathcal{A}U = VR_A, \qquad \mathcal{D}U = VR_D \tag{15}$$

for some square matrices R_A and R_D. If U_1 and V_1 are invertible, then $X = U_2U_1^{-1}$ and $Y = V_2V_1^{-1}$ satisfy (13)–(14).

If X and Y satisfy (13)–(14), then $U = \left[\begin{smallmatrix} I \\ X \end{smallmatrix}\right]$, $V = \left[\begin{smallmatrix} I \\ Y \end{smallmatrix}\right]$ satisfy (15) with $R_A = A_{11} + A_{12}X$, $R_D = D_{11} + A_{12}Y$.

If we introduce the sets

$$
\begin{aligned}
\mathbf{S}_1 &= \{(X,Y)|\ X,Y \text{ together with some } Z \text{ satisfy (8)–(10) }\} && (16)\\
\mathbf{S}_2 &= \{(X,Y)|\ X,Y \text{ satisfy (13)–(14) }\}, && (17)
\end{aligned}
$$

then $\mathbf{S}_1 \subseteq \mathbf{S}_2$ but, as the following example demonstrates, $\mathbf{S}_1 \neq \mathbf{S}_2$ in general.

Example 3.3 If

$$
\mathcal{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ \hline -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{array}\right],
$$

$$
\mathcal{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right],
$$

$$
\mathcal{C} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right],
$$

then $\mathbf{S}_1 = \emptyset$. However,

$$
\mathbf{S}_2 = \left\{ \left(\begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & x_{22} \end{bmatrix}, \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -1 - x_{22} \end{bmatrix} \right),\ x_{22} \in \mathbf{C} \right\}.
$$

The relationship between $\mathbf{S}_1$ and $\mathbf{S}_2$ is characterized in the following theorem.

Theorem 3.4 *There exist solutions X, Y and Z of matrix equations (8)–(10) if and only if there exist solutions X, Y of (13)–(14) satisfying*

$$\begin{aligned} \mathrm{kernel}(C_{11}+C_{12}X) &\subseteq \mathrm{kernel}(C_{21}+C_{22}X), \\ \mathrm{kernel}(B_{22}-YB_{12})^H &\subseteq \mathrm{kernel}(YB_{11}-B_{21})^H. \end{aligned} \tag{18}$$

Moreover,

$$\mathbf{S}_1 = \{(X,Y) |\ (X,Y) \in \mathbf{S}_2,\ X, Y \text{ satisfy } (18)\}.$$

Proof. Let $\mathcal{D} = \mathcal{BC}$. If $(X,Y) \in \mathbf{S}_2$, then (12) holds. Consider the singular value decompositions

$$B_{22} - YB_{12} = U_1 \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V_1^H, \qquad C_{11} + C_{12}X = U_2 \begin{bmatrix} \Sigma_2 & 0 \\ 0 & 0 \end{bmatrix} V_2^H,$$

where U_1, U_2, V_1, V_2 are unitary and Σ_1 and Σ_2 are nonsingular and diagonal [26]. If X, Y satisfy the conditions in (18), then there exist matrices P_{11}, P_{21}, Q_{11} and Q_{12}, such that

$$C_{21} + C_{22}X = V_1 \begin{bmatrix} P_{11} & 0 \\ P_{21} & 0 \end{bmatrix} V_2^H, \quad YB_{11} - B_{21} = U_1 \begin{bmatrix} Q_{11} & Q_{12} \\ 0 & 0 \end{bmatrix} U_2^H,$$

with $\Sigma_1 P_{11} = Q_{11}\Sigma_2$. If we set

$$Z = V_1 \begin{bmatrix} \Sigma_1^{-1}Q_{11} & \Sigma_1^{-1}Q_{21} \\ P_{21}\Sigma_2^{-1} & Z_{22} \end{bmatrix} U_2^H,$$

where Z_{22} is arbitrary, then Z satisfies

$$Z(C_{11} + C_{12}X) = C_{21} + C_{22}X, \quad (B_{22} - YB_{12})Z = YB_{11} - B_{21}, \tag{19}$$

which are just equations (9)–(10). Equations (13) and (8) are the same. Hence, X, Y and Z satisfy (8)–(10).

If X, Y and Z satisfy (8)–(10) then $(X,Y) \in \mathbf{S}_2$. Since (9)–(10) is the same system as (19), it follows that X, Y satisfy (18). $\square$

The nonsingularity of U_1, V_1 and W_1 in (7) is implicitly determined by the matrix pencil $\alpha\mathcal{A} - \beta\mathcal{BC}$, namely the coefficient matrices of the matrix equations (8)–(10) or (13)–(14). In general it is difficult to find conditions on the coefficient matrices that guarantee the invertability of U_1, V_1 and W_1, but such conditions can be derived in the special cases that we discuss below.

3.1 Algebraic Riccati equations

By choosing the blocks in matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ in particular ways we obtain important subclasses.

If we specify

$$\mathcal{B} = \begin{bmatrix} I & 0 \\ B_{21} & B_{22} \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} C_{11} & 0 \\ C_{21} & I \end{bmatrix}, \tag{20}$$

then (8)–(10) simplifies to

$$\begin{aligned} A_{21} + A_{22}X &= Y(A_{11} + A_{12}X), & (21)\\ Y &= B_{21} + B_{22}Z, & (22)\\ X &= ZC_{11} - C_{21}. & (23) \end{aligned}$$

This leads to a quadratic matrix equation in Z, which is often called *continuous-time algebraic Riccati equation*

$$A_{22}ZC_{11} - B_{22}ZA_{11} - (B_{22}Z + B_{21})A_{12}(ZC_{11} - C_{21}) + \hat{A}_{21} = 0. \tag{24}$$

Here we have set $\hat{A}_{21} = A_{21} - A_{22}C_{21} - B_{21}A_{11}$.

Corollary 3.5 *Let $\mathcal{A}$ be as in (6), $\mathcal{B}, \mathcal{C}$ as in (20) and U, V and W as in (7) and assume they satisfy (11). If W_1 is invertible, then U_1 and V_1 are invertible and $X = U_2U_1^{-1}$, $Y = V_2V_1^{-1}$ and $Z = W_2W_1^{-1}$ satisfy (21)–(24).*

Proof. Using Corollary 3.1, we only need to show that U_1, V_1 are invertible. By comparing the first block in $\mathcal{B}W = VR_B$ and considering the block diagonal structure of $\mathcal{B}$ in (20) we obtain $W_1 = V_1R_B$. Hence, if W_1 is nonsingular then V_1 is nonsingular. To prove the nonsingularity of U_1, without loss of generality we may assume that W and U have orthonormal columns i.e., $W^HW = U^HU = I$. We extend W and U to square unitary matrices

$$\mathcal{W} = \begin{bmatrix} W_1 & W_3 \\ W_2 & W_4 \end{bmatrix}, \qquad \mathcal{U} = \begin{bmatrix} U_1 & U_3 \\ U_2 & U_4 \end{bmatrix}.$$

Equation $\mathcal{C}U = WR_C$ implies that there are matrices S_C and T_C such that $\mathcal{C}\mathcal{U} = \mathcal{W}\left[\begin{smallmatrix} R_C & S_C \\ 0 & T_C \end{smallmatrix}\right]$, or equivalently

$$\mathcal{W}^H\mathcal{C} = \begin{bmatrix} R_C & S_C \\ 0 & T_C \end{bmatrix}\mathcal{U}^H. \tag{25}$$

Using the block triangular structure of $\mathcal{C}$ in (20) and comparing the $(2,2)$ blocks on both sides of (25) we get $W_4^H = T_CU_4^H$. Since $\mathcal{W}$ is unitary, using the CS decomposition [26] of $\mathcal{W}$, $\det W_1 \neq 0$ implies that $\det W_4 \neq 0$ and hence $\det U_4 \neq 0$. Since $\mathcal{U}$ is also unitary, using again the CS decomposition, we have $\det U_1 \neq 0$. ☐

For $\mathcal{B}$ and $\mathcal{C}$ as in (20), equations (13)–(14) take the form

$$A_{21} + A_{22}X = Y(A_{11} + A_{12}X), \quad B_{22}(X + C_{21}) = (Y - B_{21})C_{11}. \tag{26}$$

The existence of the solution was discussed in Corollary 3.2. Combining the results of Theorem 3.4 and Corollary 3.5 we have the following corollary.

Corollary 3.6 *In the notation of Corollary 3.5 the following are equivalent.*

(i) The matrix equation (24) has a solution.

(ii) W_1 *is nonsingular.*

(iii) There exist matrices X *and* Y *which satisfy (26) and satisfy*

$$\operatorname{kernel} C_{11} \subseteq \operatorname{kernel}(C_{21} + X), \qquad \operatorname{kernel} B_{22}^H \subseteq \operatorname{kernel}(Y - B_{21})^H.$$

If we consider the special case that $\mathcal{B} = \mathcal{C} = I$, then the eigenvalue problem is reduced to the ordinary matrix eigenvalue problem for the matrix $\mathcal{A}$ and (24) is the classical formulation of the nonsymmetric algebraic Riccati equation [13]

$$A_{22}Z - ZA_{11} - ZA_{12}Z + A_{21} = 0. \tag{27}$$

For completeness we list the relationship between deflating subspaces and solutions of (27).

Corollary 3.7 *Let* $\mathcal{A}$ *be as (6) and let* $U = \left[\begin{smallmatrix} U_1 \\ U_2 \end{smallmatrix}\right]$ *with* $U_1 \in \mathbf{C}^{n\times n}$ *such that* $\mathcal{A}U = UR_A$. *If* U_1 *is nonsingular then* $Z = U_2U_1^{-1}$ *satisfies the Riccati equation (27). Conversely, if* Z *is a solution of (27) then the columns of* $U = \left[\begin{smallmatrix} I \\ Z \end{smallmatrix}\right]$ *span an invariant subspace of* $\mathcal{A}$ *corresponding to* $\Lambda(A_{11} + A_{12}Z)$.

3.2 Symmetric algebraic Riccati equations

A special case of quadratic matrix equations that arises in optimal control theory of descriptor systems [45] is the symmetric, generalized, continuous-time algebraic Riccati equation

$$A^HZE + E^HZA - E^H(Z + F^H)D(Z + F)E + \tilde{G} = 0, \tag{28}$$

where $\tilde{G} = G + A^HF + F^HA$, $G = G^H$, $D = D^H$ and $A, D, E, F, G \in \mathbf{C}^{n\times n}$. For this equation the matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are given by

$$\mathcal{A} = \begin{bmatrix} A & -D \\ -G & -A^H \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} I & 0 \\ F^H & E^H \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} E & 0 \\ -F & I \end{bmatrix} = J^H\mathcal{B}^HJ. \tag{29}$$

The matrices $\mathcal{A}$ and $i\mathcal{BC}$ in (29) are *Hamiltonian*, i.e., $(J_n\mathcal{A})^H = J_n\mathcal{A}$ and $(J_n(i\mathcal{BC}))^H = J_n(i\mathcal{BC})$.

Equation (28) is a special case of (24). However, in practice, one is particularly interested in *Hermitian* solutions of (28). Suppose that (28) has an Hermitian solution Z. If $X = ZE + F$ and $Y = E^HZ + F^H$, then by (21)–(23) and Corollary 3.1, the matrices

$$U = \begin{bmatrix} I \\ X \end{bmatrix}, \qquad V = \begin{bmatrix} I \\ Y \end{bmatrix}, \qquad W = \begin{bmatrix} I \\ Z \end{bmatrix} \tag{30}$$

satisfy (11) with $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ from (29). Note that $Z = Z^H$ implies $X = Y^H$. This leads to the following existence result for Hermitian solutions of (28).

Theorem 3.8 *Let $\mathcal{A}$ and $\mathcal{C}$ be as in (29). If there is a Hermitian solution Z to (28), then there exist a symplectic matrix $\mathcal{W}$ (i.e., $\mathcal{W}^H J\mathcal{W} = J$), a nonsingular matrix $\mathcal{U}$ and $n \times n$ matrices R_A, S_A, R_C, S_C and T_C such that*

$$J^H \mathcal{U} J \mathcal{A} \mathcal{U} = \begin{bmatrix} R_A & S_A \\ 0 & -R_A^H \end{bmatrix}, \quad \mathcal{W}^{-1} \mathcal{C} \mathcal{U} = \begin{bmatrix} R_C & S_C \\ 0 & T_C \end{bmatrix}. \tag{31}$$

Conversely, suppose that there exist a symplectic matrix $\mathcal{W}$ and a nonsingular matrix $\mathcal{U}$ satisfying (31). Let $W = \left[\begin{smallmatrix} W_1 \\ W_2 \end{smallmatrix}\right]$ and $U = \left[\begin{smallmatrix} U_1 \\ U_2 \end{smallmatrix}\right]$ be the submatrices (with $n \times n$ blocks) formed from the first n columns of $\mathcal{W}$ and $\mathcal{U}$, respectively. If W_1 is nonsingular then U_1 is also nonsingular and $Z = W_2 W_1^{-1}$ is an Hermitian solution of (28).

Proof. Let Z be an Hermitian solution of (28) and let $X = ZE + F$, $Y = E^H Z + F^H = X^H$. Defining U, V and W as in (30), by Corollary 3.1, U, V and W satisfy (11) with $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ defined in (29) and $R_A = A - DZ$, $R_B = I$, $R_C = E$. Introducing

$$\mathcal{W} = \begin{bmatrix} I & 0 \\ Z & I \end{bmatrix}, \quad \mathcal{U} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} I & 0 \\ Y & I \end{bmatrix},$$

we have $\mathcal{W}^H J \mathcal{W} = J$ (because $Z = Z^H$), i.e., $\mathcal{W}$ is symplectic. Furthermore, $\mathcal{V}^{-1} = J^H \mathcal{U}^H J$. From (11) we have (31).

If (31) is satisfied, then we have

$$J^H \mathcal{U}^H J \mathcal{B} \mathcal{W} = \begin{bmatrix} T_C^H & -S_C^H \\ 0 & R_C^H \end{bmatrix}$$

and by Corollary 3.5, $Z = W_2 W_1^{-1}$ satisfies (28). Since $\mathcal{W}$ is symplectic, Z is Hermitian. □

If we are not interested in the solution Z but rather in the matrices X or Y [45], then we may restrict ourselves to the pair of matrix equations

$$A^H X + YA - YDX + G = 0, \qquad E^H (X - F) = (Y - F^H) E. \tag{32}$$

The related matrix pencil is $\alpha\mathcal{A} - \beta\mathcal{D}$ with

$$\mathcal{A} = \begin{bmatrix} A & -D \\ -G & -A^H \end{bmatrix}, \qquad \mathcal{D} = \begin{bmatrix} E & 0 \\ F^H - F & E^H \end{bmatrix}. \tag{33}$$

Here $\mathcal{A}$ and $i\mathcal{D}$ are Hamiltonian. For the analysis of such pencils see [42, 43] and for numerical methods for the computation of deflating subspaces for such matrices see [9, 10].

The solvability condition for (32) was given in Corollary 3.2. The solution set is just $\mathbf{S}_2$ defined in (17). We also can define a set $\mathbf{S}_1^H$ analogous to $\mathbf{S}_1$ as in (16), but with the further restriction that Z is Hermitian. Moreover, we introduce a third set as

$$\mathbf{S}_3 = \{(X, Y) | (X, Y) \in \mathbf{S}_2, \; X = Y^H\}.$$

For the solutions in $\mathbf{S}_3$ we have the following theorem.

Theorem 3.9 *Consider the matrix pencil $\alpha\mathcal{A}-\beta\mathcal{D}$ defined via (33). If there exist solutions X and Y of (32) with $X = Y^H$, then there exists a nonsingular matrix $\mathcal{U} \in \mathbf{C}^{2n\times 2n}$ and $n \times n$ matrices R_A, S_A, R_D, and S_D such that*

$$J^H\mathcal{U}^H J\mathcal{A}\mathcal{U} = \begin{bmatrix} R_A & S_A \\ 0 & -R_A^H \end{bmatrix}, \qquad J^H\mathcal{U}^H J\mathcal{D}\mathcal{U} = \begin{bmatrix} R_D & S_D \\ 0 & R_D^H \end{bmatrix}. \tag{34}$$

If such a matrix $\mathcal{U}$ exists, let $U = \left[\begin{smallmatrix} U_1 \\ U_2 \end{smallmatrix}\right]$ (with $n \times n$ blocks) be the submatrix composed of the first n columns of $\mathcal{U}$. If U_1 is invertible, then $X = U_2U_1^{-1}$ and $Y = X^H$ satisfy (32). Moreover, if (X, Y) satisfy (32) and $\operatorname{kernel} E = \operatorname{kernel}(X - F)$, *then (28) has an Hermitian solution.*

Proof. The proof follows directly from Theorem 3.8 and Corollary 3.6. □

Clearly we have $\mathbf{S}_1^H \subseteq \mathbf{S}_3 \subseteq \mathbf{S}_2$ but in general the inclusions are strict as the following examples demonstrate.

Example 3.10 If

$$A = G = I_2, \qquad F = 0, \qquad E = D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

then $\mathbf{S}_1 = \mathbf{S}_1^H = \emptyset$. However,

$$\mathbf{S}_2 = \left\{ \left(\begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & x_{22} \end{bmatrix}, \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -1 - x_{22} \end{bmatrix} \right),\ x_{22} \in \mathbf{C} \right\},$$

and

$$\mathbf{S}_3 = \left\{ \left(\begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -\frac{1}{2} + ia \end{bmatrix} \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -\frac{1}{2} - ia \end{bmatrix} \right),\ a \in \mathbf{R} \right\}.$$

Example 3.11 *Let $G = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and let A, D, E and F be as in Example 3.10. In this case (28) has Hermitian solutions $Z = \left[\begin{smallmatrix} 1\pm\sqrt{2} & 0 \\ 0 & z \end{smallmatrix}\right]$, $z \in \mathbf{R}$, and we have the following solution sets.*

$$\begin{aligned}
\mathbf{S}_1^H &= \{(X,Y) | X = Y = \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}\}, \\
\mathbf{S}_2 &= \{(\begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & x_{22} \end{bmatrix}, \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -x_{22} \end{bmatrix}),\ x_{22} \in \mathbf{C}\}, \\
\mathbf{S}_3 &= \{(\begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & ia \end{bmatrix} \begin{bmatrix} 1 \pm \sqrt{2} & 0 \\ 0 & -ia \end{bmatrix}),\ a \in \mathbf{R}\}.
\end{aligned}$$

We see from Example 3.11 that if there exist Hermitian solutions of (28), then using the right deflating subspace of the matrix pencil $\alpha\mathcal{A} - \beta\mathcal{D}$ in (33) to compute X may not yield the desired result. If E is nonsingular, then $\mathbf{S}_1^H = \mathbf{S}_3$, and if $(X, Y) \in \mathbf{S}_3$ then

$Z = (X - F)E^{-1}$ is an Hermitian solution of (28). But this relation does not hold in general if E is singular, see also [45].

An even more special case is the classical continuous-time algebraic Riccati equation,

$$A^H Z + ZA - ZDZ + G = 0, \tag{35}$$

which is the case that in (28) we have $E = I$ and $F = 0$. Here, the pencil is just $\mathcal{A} - \lambda I$ with the Hamiltonian matrix $\mathcal{A}$ defined in (29). From Theorem 3.8, we have the following well-known corollary, see, e.g., [38, 45].

Corollary 3.12 *Let $\mathcal{A}$ be as in (29). Suppose there exists a symplectic matrix $\mathcal{W}$ such that*

$$\mathcal{W}^{-1}\mathcal{A}\mathcal{W} = \begin{bmatrix} R_A & S_A \\ 0 & -R_A^H \end{bmatrix} \tag{36}$$

with $n \times n$ blocks R_A and S_A. Let $W = \left[\begin{smallmatrix} W_1 \\ W_2 \end{smallmatrix}\right]$ (with $n \times n$ blocks) be composed of the first n columns of $\mathcal{W}$. If W_1 is nonsingular, then $Z = W_2 W_1^{-1}$ is an Hermitian solution of (35).

The triangular forms (31), (34) and (36) do not always exist. Necessary and sufficient conditions for the existence of such triangular forms were recently given in [43, 46]. But as we have seen, even if these triangular forms exist, the existence of Hermitian solutions of (28) and (35) is not guaranteed. Several conditions which partially characterize the existence of solutions are known, see [38, 45, 53].

3.3 Matrix sign function, disc function and matrix square roots

Quadratic matrix equations include as special cases matrix square roots. Consider the matrices

$$\mathcal{A} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix}, \qquad \mathcal{B} = \begin{bmatrix} I & 0 \\ 0 & B_{22} \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} C_{11} & 0 \\ 0 & I \end{bmatrix}$$

in (20). The related matrix equation (24) then has the form

$$B_{22} Z A_{12} Z C_{11} = A_{21}. \tag{37}$$

In the more special case that $\mathcal{B} = \mathcal{C} = I$, (37) is just related to the invariant subspace problem for the matrix $\mathcal{A}$. If, furthermore, $A_{12} = I$ and $A_{21} = A$, then (37) is

$$Z^2 = A.$$

So in this case any solution Z is just a square root of A. Existence conditions for the matrix square root are discussed in [31]. In view of the relationship to invariant subspaces we have the following corollary.

Corollary 3.13 *Let $\mathcal{A} = \left[\begin{smallmatrix} 0 & I_n \\ A & 0 \end{smallmatrix}\right]$ with $A \in \mathbf{C}^{n\times n}$ and let $R \in \mathbf{C}^{n\times n}$ and $W = \left[\begin{smallmatrix} W_1 \\ W_2 \end{smallmatrix}\right] \in \mathbf{C}^{2n\times n}$ (with $n \times n$ blocks W_1 and W_2) be such that $\mathcal{A}W = WR$. If W_1 is nonsingular, then $Z = W_2 W_1^{-1} = W_1 R W_1^{-1}$ is a square root of A.*

Another important special case is that $A_{12} = A_{21} = A \in \mathbf{C}^{n\times n}$. In this case, (37) reduces to

$$ZAZ = A. \tag{38}$$

By properly choosing the invariant subspace we obtain the matrix sign function [33, 49] which is the $m = 2$ sector case of the matrix sector function.

Definition 3.14 Given a positive integer $m \geq 2$ we may partition the complex plane into m sectors

$$\Omega_k(m) = \{re^{\theta i} | \frac{(2k-3)\pi}{m} < \theta < \frac{(2k-1)\pi}{m}, r > 0\}, \qquad k = 1, \ldots, m.$$

1. A matrix Z is called an *m-th root* of a square matrix A if $Z^m = A$, and Z is called the *principal m-th root* if Z is an *m-th root* and if $\Lambda(Z) \subset \Omega_1(m)$.
2. If A^m has a principal m-th root Z, then the matrix $S := Z^{-1}A$ is called the *m-th sector function* of A.

For the matrix sign function we obtain the following corollary.

Corollary 3.15 *Suppose that $A \in \mathbf{C}^{n\times n}$ has no purely imaginary eigenvalues. Let $\mathcal{A} = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$. Let $W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \in \mathbf{C}^{2n\times n}$ (with $n\times n$ blocks W_1 and W_2) be such that $\mathcal{A}W = WR$ with all eigenvalues of R in the open right half plane. Then W_1 is nonsingular and $Z = W_2W_1^{-1}$ is the sign function of A. Moreover, $S = AZ = W_1RW_1^{-1}$ is the principal square root of A^2.*

Proof. See [11]. □

Note that the matrix sign function is only one of the solutions of (38). Different solutions are related to different invariant subspaces of $\mathcal{A}$ corresponding to different R. Note also that (38) always has the solution I.

The disc function [7, 8, 49] of a matrix A is defined through its ordered Jordan form $A = T\begin{bmatrix} J_0 & 0 \\ 0 & J_\infty \end{bmatrix}T^{-1}$, where $J_0 \in \mathbf{C}^{k\times k}$ contains the Jordan blocks corresponding to eigenvalues inside the unit disc and $J_\infty \in \mathbf{C}^{n-k\times n-k}$ corresponds to eigenvalues outside the unit disc. Then the *matrix disc function of* A is

$$\mathrm{disc}(A) = T\begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}T^{-1}.$$

(If A has an eigenvalue of modulus 1, then the disc function is undefined.) It can be shown [7] that

$$\mathrm{disc}(A) = \frac{1}{2}\left(I - \mathrm{sign}\left((A-I)^{-1}(A+I)\right)\right).$$

The disc function $D = \mathrm{disc}(A)$ is related to the deflating subspace of

$$\alpha\mathcal{A} - \beta\mathcal{D} = \alpha\begin{bmatrix} 0 & A \\ I & -I \end{bmatrix} - \beta\begin{bmatrix} 0 & I \\ A & -A \end{bmatrix}$$

corresponding to the eigenvalues inside the unit disc via

$$\begin{bmatrix} 0 & A \\ I & -I \end{bmatrix} \begin{bmatrix} I \\ D \end{bmatrix} = \begin{bmatrix} 0 & I \\ A & -A \end{bmatrix} \begin{bmatrix} I \\ D \end{bmatrix} R, \tag{39}$$

where the eigenvalues of R are the eigenvalues of $\alpha\mathcal{A} - \beta\mathcal{D}$ that lie inside the unit disc [8]. The eigenvalues of R are the union of the eigenvalues of A inside the unit disc and the reciprocals of the eigenvalues of A outside the unit disc. In particular, there are exactly n eigenvalues of $\alpha\mathcal{A} - \beta\mathcal{D}$ inside the unit disc along with a corresponding n-dimensional deflating subspace spanned by the columns of range $\left[\begin{smallmatrix} I \\ D \end{smallmatrix}\right]$. The deflating subspace spanned by the columns of $\left[\begin{smallmatrix} I \\ D \end{smallmatrix}\right]$ and therefore D are uniquely defined.

In order to derive a corresponding matrix equation via Proposition 2.2 or Corollary 3.2, we need a deflating subspace relation of the form $\mathcal{A}U = VR_A$ and $\mathcal{D}U = VR_B$, where $U = \left[\begin{smallmatrix} U_1 \\ U_2 \end{smallmatrix}\right]$, $V = \left[\begin{smallmatrix} V_1 \\ V_2 \end{smallmatrix}\right]$ and the columns of U span the deflating subspace corresponding to eigenvalues inside the unit disc. From (39) we get that U_1 is nonsingular and $D = U_2U_1^{-1}$. However, by Corollary 3.2, V_1 nonsingular would imply that $I - D - V_2V_1^{-1}AD = 0$. If the spectrum of A is not contained in the open unit disc, D is singular and the latter equation leads to a contradiction. This shows that for the matrix disc function the relation between deflating subspaces and matrix equations is not as obvious as for the matrix sign function.

If we assume that A is nonsingular, then $\alpha\mathcal{A} - \beta\mathcal{D}$ is equivalent to the matrix pencil

$$\alpha\tilde{\mathcal{A}} - \beta\tilde{\mathcal{D}} = \alpha \begin{bmatrix} I & A^2 - I \\ 0 & A^2 \end{bmatrix} - \beta \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix}$$

and the matrix disc function of A satisfies

$$\begin{aligned} \begin{bmatrix} I & A^2 - I \\ 0 & A^2 \end{bmatrix} \begin{bmatrix} I \\ D \end{bmatrix} &= \begin{bmatrix} I \\ D \end{bmatrix} (I + (A^2 - I)D) \\ \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} \begin{bmatrix} I \\ D \end{bmatrix} &= \begin{bmatrix} I \\ D \end{bmatrix} A. \end{aligned}$$

Hence, if A is nonsingular, then range $\left[\begin{smallmatrix} I \\ D \end{smallmatrix}\right]$ is the deflating subspace of $\alpha\tilde{\mathcal{A}} - \beta\tilde{\mathcal{D}}$ corresponding to eigenvalues inside the unit disc. The associated matrix equation is

$$ADA = DA^2D \tag{40}$$

and the disc function is the root of (40) for which $AD = DA$ and the nonzero eigenvalues of AD consist of the eigenvalues of A that lie inside the unit disc.

Equation (40) is satisfied by $D = \text{disc}(A)$ even when A is singular. However, if $Ax = 0$, $x \neq 0$ and D is the disc function, then $D + xx^H$ is also a root and $AD = A(D + xx^H)$. Hence, in this case, the matrix AD does not distinguish $D = \text{disc}(A)$ from other roots of (40). Also, if A is singular, then $\alpha\tilde{\mathcal{A}} - \beta\tilde{\mathcal{D}}$ is not regular and the deflating subspace is no longer uniquely defined.

4 Rational matrix equations

Analogous to the construction of continuous-time algebraic Riccati equations, the corresponding discrete-time equations also arise as special cases.

4.1 Algebraic Riccati equations

The case $m = 2$ in Proposition 2.3 also leads to some classical rational matrix equations. If we set

$$\mathcal{A} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} C_{11} & 0 \\ 0 & I \end{bmatrix}, \tag{41}$$

in (6), then the equations in (7) become

$$A_{21} + A_{22}X = YA_{11}, \quad B_{22}Z = Y(B_{11} + B_{12}Z), \quad X = ZC_{11}. \tag{42}$$

We then obtain a rational matrix equation for Z, the *discrete-time algebraic Riccati equation* as

$$A_{22}ZC_{11} - B_{22}Z(B_{11} + B_{12}Z)^{-1}A_{11} + A_{21} = 0,$$

or equivalently

$$A_{22}ZC_{11} - B_{22}ZA_{11} + B_{22}Z(B_{11} + B_{12}Z)^{-1}(B_{12}Z + B_{11} - I)A_{11} + A_{21} = 0. \tag{43}$$

The existence of solutions for (43) as well as (42) follows from Corollary 3.1 with the matrices in (41) but with an additional restriction for the nonsingularity of $B_{11} + B_{12}Z$. Another formulation, using generalized inverses allows to drop this condition [1].

Theorem 4.1 *Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ be as in (41) and let U, V and W be as in (7) satisfying (11). If W_1 and $B_{11}W_1 + B_{12}W_2$ are invertible, then U_1 and V_1 are invertible and $X = U_2U_1^{-1}$, $Y = V_2V_1^{-1}$ and $Z = W_2W_1^{-1}$ satisfy (42) and (43).*

Proof. The proof is similar to that of Corollary 3.7. □

4.2 Symmetric discrete-time algebraic Riccati equations

Analogous to the continuous-time case we also have the symmetric cases. The symmetric form of (43) is the generalized, symmetric, discrete-time algebraic Riccati equation

$$E^HZE - A^HZA + A^HZ(I + DZ)^{-1}DZA + G = 0, \qquad D = D^H, \quad G = G^H, \tag{44}$$

with the corresponding matrices

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ G & E^H \end{bmatrix}, \qquad \mathcal{B} = \begin{bmatrix} I & D \\ 0 & A^H \end{bmatrix}, \qquad \mathcal{C} = \begin{bmatrix} E & 0 \\ 0 & I \end{bmatrix}. \tag{45}$$

Analogous to Theorem 3.8 we have the following existence and uniqueness result.

Theorem 4.2 *Let $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ be as in (45). If there exists a symplectic matrix $\mathcal{W}$ and nonsingular matrices $\mathcal{U}$ and $\mathcal{V}$ such that*

$$\mathcal{V}\mathcal{A}\mathcal{U} = \begin{bmatrix} R_A & S_A \\ 0 & T_A \end{bmatrix}, \quad \mathcal{V}\mathcal{B}\mathcal{W} = \begin{bmatrix} R_B & S_B \\ 0 & T_B \end{bmatrix}, \quad \mathcal{W}^{-1}\mathcal{C}\mathcal{U} = \begin{bmatrix} R_C & S_C \\ 0 & T_C \end{bmatrix}, \tag{46}$$

with $n \times n$ blocks R_A, S_A, T_A, R_B, S_B, T_B, R_C, S_C and T_C, then there exists an Hermitian solution of (44).

Suppose that $\mathcal{W}$, $\mathcal{U}$ and $\mathcal{V}$ satisfy (46) and that $\mathcal{W}$ is symplectic. Let $W = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix}$, $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$, $V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$ (with $n \times n$ blocks W_i, U_i and V_i) be the submatrices formed from the first n columns of $\mathcal{W}$, $\mathcal{U}$ and $\mathcal{V}$, respectively. If W_1 and $W_1 + DW_2$ are nonsingular, then U_1 and V_1 are also nonsingular and $Z = W_2 W_1^{-1}$ is an Hermitian solution of (44).

Proof. The proof is analogous to that of Theorem 3.8. □

In practice, see [45], one often needs the solution $X = ZE$ rather than Z. This solution can be obtained by computing a proper right deflating subspace of the pencil $\alpha \begin{bmatrix} A & 0 \\ G & E^H \end{bmatrix} - \beta \begin{bmatrix} E & D \\ 0 & A^H \end{bmatrix}$. However, as in the continuous-time case this subspace is guaranteed to give the desired solution only if E is nonsingular.

5 Linear matrix equations

The nonlinear part in the matrix equations (8)–(10) and (13)–(14) is contributed by the $(1,2)$ blocks of the matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D} = \mathcal{B}\mathcal{C}$. If all the $(1,2)$ blocks are zero, then (8)–(10) reduce to

$$\begin{aligned} A_{22}X - YA_{11} + A_{21} &= 0, \\ B_{22}Z - YB_{11} + B_{21} &= 0, \\ C_{22}X - ZC_{11} + C_{21} &= 0, \end{aligned} \tag{47}$$

and (13)–(14) reduce to

$$A_{22}X - YA_{11} + A_{21} = 0, \quad D_{22}X - YD_{11} + D_{21} = 0, \tag{48}$$

respectively. In the nonlinear case, the eigenstructure of $\alpha R_A - \beta R_B R_C$ or $\alpha R_A - \beta R_D$, which corresponds to the deflating subspaces, may be nonunique. This implies that different solutions related to different eigenstructures may exist. In the linear case, however, the eigenstructure is essentially fixed. This can be easily observed from (11) and (15), since if the solutions exist, then $\alpha R_A - \beta R_B R_C$ and $\alpha R_A - \beta R_D$ are equivalent (pencil equivalent) to $\alpha A_{11} - \beta B_{11} C_{11}$ and $\alpha A_{11} - \beta D_{11}$, respectively.

The linear matrix equations have been studied extensively, [20, 39, 50, 54]. Here we will briefly discuss the existence problem for equations (47) and (48). Since they are just special cases of the nonlinear equations, all results in the previous sections still apply. On the other hand because of the linearity, the conditions can be described in a way that is more directly related to the matrices $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$.

The condition for the existence and uniqueness of the solutions X and Y of (48) can be stated as follows.

Corollary 5.1 *Consider the matrices* $\mathcal{A} = \left[\begin{smallmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{smallmatrix}\right]$, $\mathcal{D} = \left[\begin{smallmatrix} D_{11} & 0 \\ D_{21} & D_{22} \end{smallmatrix}\right]$ *as well as*

$$\hat{\mathcal{A}} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \quad \hat{\mathcal{D}} = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix}.$$

The matrix equation (48) has a solution if and only if $\alpha\mathcal{A} - \beta\mathcal{D}$ *is pencil equivalent to* $\alpha\hat{\mathcal{A}} - \beta\hat{\mathcal{D}}$. *There is a unique solution if and only if* $\alpha\mathcal{A} - \beta\mathcal{D}$ *is regular and* $\Lambda(A_{11}, D_{11}) \cap \Lambda(A_{22}, D_{22}) = \emptyset$.

Proof. See [20] and [54]. □

For completeness, in the remainder of this subsection we list the linear matrix equations with a single unknown matrix and the related matrix pencils. The existence and uniqueness of the solution can be derived by combining the results in the previous subsections with Corollary 5.1.

Generalized Sylvester equations have the form

$$A_{22}ZC_{11} - B_{22}ZA_{11} + \tilde{A}_{21} = 0$$

where $\tilde{A}_{21} = A_{21} - A_{22}C_{21} - B_{21}A_{11}$ and the related pencil is

$$\mathcal{A} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} I & 0 \\ B_{21} & B_{22} \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} C_{11} & 0 \\ C_{21} & I \end{bmatrix}. \tag{49}$$

The results in Corollary 3.5 can be applied to this equation. Note that the deflating subspace must correspond to $\alpha A_{11} - \beta C_{11}$. With $\mathcal{D} = \mathcal{B}\mathcal{C}$ we can combine the results in Corollary 3.6 and Corollary 5.1 to get necessary and sufficient conditions for the existence of solutions.

Generalized Lyapunov equations have the form

$$A^H ZE + E^H ZA + \tilde{G} = 0,$$

where $\tilde{G} = G + A^H F + F^H A$, $G = G^H$, $D = D^H$ and $A, E, F, G \in \mathbf{C}^{n\times n}$. The related matrix pencil is

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ -G & -A^H \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} I & 0 \\ F^H & E^H \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} E & 0 \\ -F & I \end{bmatrix} = J^H \mathcal{B}^H J.$$

For such equations we can apply Theorems 3.8, 3.9 and Corollary 5.1.

Generalized Stein equations have the form

$$E^H ZE - A^H ZA + G = 0, \quad G = G^H,$$

with

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ G & E^H \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} I & 0 \\ 0 & A^H \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} E & 0 \\ 0 & I \end{bmatrix}.$$

This is the linear version of the symmetric discrete-time algebraic Riccati equation.

The classical Sylvester equation is

$$A_{22}Z - ZA_{11} + A_{21} = 0, \tag{50}$$

with

$$\mathcal{A} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \quad \mathcal{B} = \mathcal{C} = I \tag{51}$$

and the classical Lyapunov equation is

$$A^H Z + ZA + G = 0, \quad G = G^H,$$

with

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ -G & -A^H \end{bmatrix}, \quad \mathcal{B} = \mathcal{C} = I.$$

Finally, the Stein equation is

$$Z - A^H ZA + G = 0, \quad G = G^H,$$

with

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ G & I \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} I & 0 \\ 0 & A^H \end{bmatrix}, \quad \mathcal{C} = I.$$

Remark 5.2 The discussed relationship between deflating subspaces and matrix equations can be extended to more general matrix equations. For instance we may consider the linear matrix equations [20]

$$AXB + CYD = E, \quad GXH + KYL = F.$$

However, the general linear matrix equation

$$\sum_{k=0}^{m} A_k Z B_k = 0,$$

[36, 37] does not appear to have a related deflating subspace.

6 Numerical methods for $m = 2$

Since the periodic QZ decomposition can be computed by applying the periodic QZ algorithm [15, 29], in principle all deflating and/or invariant subspaces discussed in this paper can be computed in a numerically stable way. We will call a method based on this approach a *subspace method*. For matrix pencils with matrices as in (6), we may directly apply the periodic QZ algorithm. In some special cases, however, the periodic QZ algorithm may be modified to adapt to the special structure. Much can be gained from exploiting the structure of the symmetric equations (28), (35) and (44). Theorems 3.8 and 4.2 and Corollary 3.12 show that for these symmetric equations the related matrix pencils have certain symmetry structures. Special equivalence transformations may be employed to preserve these structures, see [2, 11, 12, 18, 19, 44, 45]. However, a numerically stable and efficient method for computing the structured decompositions (31), (34), (36) and (46) in general is still an open problem.

For the eigenvalue problem corresponding to (37) there is a simplified QR like method for computing the generalized Schur form if A_{21} and A_{12} are square.

The numerical methods for linear matrix equations can be simplified by using the block triangular forms of the related matrices and the properties of the related eigenvalues. Taking the generalized Sylvester equation as an example, where the matrices are as in (49), we obtain a periodic QR-like method as follows.

First we compute the generalized Schur forms of the matrix pencils $\alpha A_{11} - \beta C_{11}$ and $\alpha A_{22} - \beta B_{22}$ respectively. Then we apply the eigenvalue reordering method [26], to the block lower triangular pencil to transform the pencil as

$$\alpha \begin{bmatrix} \hat{A}_{11} & 0 \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} - \beta \begin{bmatrix} \hat{B}_{11} & 0 \\ \hat{B}_{21} & \hat{B}_{22} \end{bmatrix} \begin{bmatrix} \hat{C}_{11} & 0 \\ \hat{C}_{21} & \hat{C}_{22} \end{bmatrix},$$

with $\alpha \hat{A}_{22} - \beta \hat{B}_{22} \hat{C}_{22}$ equivalent to $\alpha A_{11} - \beta C_{11}$. By exchanging block rows and columns simultaneously the matrix pencil is finally equivalent to

$$\alpha \begin{bmatrix} \hat{A}_{22} & \hat{A}_{21} \\ 0 & \hat{A}_{11} \end{bmatrix} - \beta \begin{bmatrix} \hat{B}_{22} & \hat{B}_{21} \\ 0 & \hat{B}_{11} \end{bmatrix} \begin{bmatrix} \hat{C}_{22} & \hat{C}_{21} \\ 0 & \hat{C}_{11} \end{bmatrix}.$$

The desired interior deflating subspace can be read off from this form.

Many efficient numerical algorithms have already been designed for computing the solutions of special linear and nonlinear matrix equations. For the case of linear equations the basic algorithm was given in [6] and the generalized in [20]. For matrix square roots there are similar methods in [14, 30]. We call these methods *direct methods*. The direct method implicitly computes a basis of the invariant or deflating subspace as $\left[\begin{smallmatrix} I \\ Z \end{smallmatrix}\right]$ with a solution Z. (In practice only Z is computed.) So the difference between direct and subspace methods is that in the latter an orthonormal basis for the subspace is computed.

The above analysis shows that often deflating subspace and the solution of the related matrix equation can be computed from each other. This fact is widely used in practice.

For example the Sylvester equation is used for Jordan canonical form reduction [26], and the invariant subspace method is used for the solution of Riccati equations [4, 13, 51, 45]. However, in finite precision arithmetic two mathematically equivalent methods may give quite different results. In order to point out the difficulties that may arise, we study, as an example, the Sylvester equation (50) which is related to the invariant subspace problem for the matrix $\mathcal{A}$ given in (51). Assume that $\Lambda(A_{11}) \cap \Lambda(A_{22}) = \emptyset$, so that (50) has a unique solution Z. Let $\mathcal{U} = \left[\begin{smallmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{smallmatrix}\right]$ be unitary such that

$$\mathcal{U}^H \mathcal{A} \mathcal{U} = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} = \mathcal{R}, \quad \Lambda(R_{11}) = \Lambda(A_{11}). \tag{52}$$

Since $\mathcal{U}$ is unitary, we have that

$$Z = U_{21} U_{11}^{-1} = -U_{22}^{-H} U_{12}^H. \tag{53}$$

Denote by $\sigma_{\min}(A)$ the minimum singular value of the matrix A. Using (53) and the orthonormality of $\left[\begin{smallmatrix} U_{11} \\ U_{21} \end{smallmatrix}\right]$ we have [34]

$$\|U_{11}^{-1}\|_2 = \sqrt{1 + \|Z\|_2^2}, \qquad \|U_{11}\|_2 = \sqrt{\frac{1}{1 + \sigma_{\min}(Z)^2}}. \tag{54}$$

Let ε be a small number of the order of the roundoff unit and let $\mathcal{U}_s$ and $\mathcal{R}_s$ be the matrices in (52), computed by a backward stable numerical method. Then there exists a matrix $\mathcal{E}$, with $\|\mathcal{E}\|_2 \leq \gamma_1 \varepsilon \|\mathcal{A}\|_2$, such that

$$\mathcal{U}_s^H (\mathcal{A} + \mathcal{E}) \mathcal{U}_s = \mathcal{R}_s.$$

Let $\mathcal{U}_s$ be partitioned conformally with $\mathcal{U}$ as $\mathcal{U}_s := [\hat{U}_1, \hat{U}_2] := \begin{bmatrix} \hat{U}_{11} & \hat{U}_{12} \\ \hat{U}_{21} & \hat{U}_{22} \end{bmatrix}$, and set

$$\mathcal{E}_s = \hat{U}_2^H \mathcal{A} \hat{U}_1 = -\hat{U}_2^H \mathcal{E} \hat{U}_1. \tag{55}$$

Then

$$\|\mathcal{E}_s\|_2 \leq \gamma_1 \varepsilon \|\mathcal{A}\|_2, \tag{56}$$

which can be viewed as the residual of the problem of computing the invariant subspace range $\hat{U}_1$.

Let Z_d be the solution of equation (50) computed with a backward stable numerical method and let $\mathcal{F}_d = A_{22} Z_d - Z_d A_{11} + A_{21}$ be the corresponding residual, then from [25] we obtain

$$\|\mathcal{F}_d\|_2 = \|A_{22} Z_d - Z_d A_{11} + A_{21}\|_2 \leq \gamma_2 \varepsilon \|\mathcal{A}\|_2 \|Z\|_2. \tag{57}$$

If our primary goal is to compute Z and if we use the subspace method, then let Z_s be the matrix computed as $\hat{U}_{21} \hat{U}_{11}^{-1}$ with corresponding residual $\mathcal{F}_s = A_{22} Z_s - Z_s A_{11} + A_{21}$. By using (55), (56), (53) and (54) we have

$$\|\mathcal{F}_s\|_2 = \|A_{22} Z_s - Z_s A_{11} + A_{21}\|_2 \leq \gamma_3 \varepsilon \|\mathcal{A}\|_2 (1 + \|Z\|_2^2). \tag{58}$$

Inequalities (57) and (58) suggest that the solution computed via the subspace method may sometimes be less inaccurate than the solution obtained via a direct method. A Riccati equation example in which this happens in actual computation appears in [48].

On the other hand if the primary goal is to compute an orthonormal basis of the invariant subspace corresponding to $\Lambda(A_{11})$ and we use a direct method, let $\mathcal{U}_d$ be the computed unitary matrix such that

$$\begin{bmatrix} I \\ Z_d \end{bmatrix} + E_Z = \mathcal{U}_d \begin{bmatrix} T \\ 0 \end{bmatrix}, \quad \|E_Z\|_2 \le \gamma_4 \varepsilon \sqrt{\|Z\|_2^2 + 1}$$

which is a QR decomposition. Denote by E_d the $(2,1)$ block of $\mathcal{U}_d^H \mathcal{A} \mathcal{U}_d$. Then by (57) and the perturbation theory for the QR decomposition [26] we have

$$\|E_d\|_2 \le \gamma_5 \varepsilon \frac{\|\mathcal{A}\|_2 \sqrt{1 + \|Z\|_2^2}}{\sqrt{1 + \sigma_{\min}(Z)^2}}. \tag{59}$$

Inequalities (56) and (59) suggest that the subspace method may sometimes yield better results than the direct method.

The significance of the orthonormal basis is indicated in the following example. Consider the problem of computing the Jordan canonical form of a square matrix A. Suppose that we have already determined the Schur form of $A+E$ with E a small perturbation (say, using the QR algorithm), i.e., we have determined a unitary matrix Q and (for convenience) a lower triangular matrix R such that

$$Q^H (A+E) Q = R =: \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix},$$

where we assume that $\Lambda(R_{11}) \cap \Lambda(R_{22}) = \emptyset$. To extract further information about the Jordan canonical form, further reductions, see [26], are carried out by removing first the block R_{21}. To do this a matrix $X = \left[\begin{smallmatrix} I & 0 \\ Z & I \end{smallmatrix}\right]$ is determined so that

$$R_1 := \begin{bmatrix} R_{11} & 0 \\ 0 & R_{22} \end{bmatrix} = (QX)^{-1}(A+E)(QX).$$

Here the matrix Z satisfies the Sylvester equation $R_{22}Z - ZR_{11} + R_{21} = 0$. Clearly the first n columns of X span an n-dimensional invariant subspace of $A + E$.

On the other hand let $Y = [Y_1, Y_2] = \left[\begin{smallmatrix} G & 0 \\ ZG & I \end{smallmatrix}\right]$, with $G = (I + Z^H Z)^{-\frac{1}{2}}$, where $F^{\frac{1}{2}}$ denotes the unique positive definite square root of the positive definite matrix F. Then Y_1 forms an orthonormal basis of $\left[\begin{smallmatrix} I \\ Z \end{smallmatrix}\right]$ and one can easily verify that

$$(QY)^{-1}(A+E)(QY) = \begin{bmatrix} G^{-1}R_{11}G & 0 \\ 0 & R_{22} \end{bmatrix} =: R_2.$$

Both $(QX)^{-1}A(QX)$ and $(QY)^{-1}A(QY)$ are similar to A. If we set $E_1 = (QX)^{-1}E(QX)$ and $E_2 = (QY)^{-1}E(QY)$, then R_1 has a distance to a matrix which is similar to A measured by $\|E_1\|_2$, and R_2 has a distance measured by $\|E_2\|_2$. Note that

$$\begin{aligned} \|X\|_2 &= \|X^{-1}\|_2 = \frac{1}{2}(\|Z\|_2 + \sqrt{\|Z\|_2^2 + 4}), \\ \|Y\|_2 &= \sqrt{1 + \frac{\|Z\|_2}{\sqrt{1+\|Z\|_2^2}}}, \\ \|Y^{-1}\|_2 &= \sqrt{\|Z\|_2^2 + 1 + \|Z\|_2\sqrt{\|Z\|_2^2+1}} \end{aligned}$$

and hence

$$\|E_1\|_2 \leq \frac{1}{2}(\|Z\|_2^2 + 2 + \|Z\|_2\sqrt{\|Z\|_2^2+4})\|E\|_2, \quad \|E_2\|_2 \leq (\|Z\|_2 + \sqrt{1+\|Z\|_2^2})\|E\|_2.$$

If $\|Z\|_2$ is large, then $\|E_1\|_2$ may be much larger than $\|E_2\|_2$ by a factor $\|Z\|_2$. This suggests that R_2 may give more precise information about the Jordan structure than R_1.

7 Polynomial systems

By choosing $m > 2$ in Proposition 2.3 we can derive higher order polynomial or rational matrix equations. We will focus here on m-th roots of matrices.

To do this we specify the matrices $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ in Proposition 2.3 as

$$\mathcal{A} = \begin{bmatrix} 0 & A_{12} & & \\ & \ddots & \ddots & \\ & & \ddots & A_{m-1,m} \\ A_{m,1} & & & 0 \end{bmatrix}, \quad \mathcal{B} = \mathcal{C} = I, \tag{60}$$

with $m \geq 3$. This leads to an eigenvalue problem for the $m \times m$ block matrix $\mathcal{A}$. The equations in (2)–(4) become

$$A_{m,1} = Z_{m-1}A_{12}Z_1, \tag{61}$$

$$A_{k,k+1}Z_k = Z_{k-1}A_{12}Z_1, \qquad k = 2, \ldots, m-1, \tag{62}$$

Multiplying $A_{2,3}\cdots A_{m-1,m}$ from the left to the last equation, using the other $m-2$ equations, we get

$$(Z_1A_{12})^{m-1}Z_1 = (\prod_{k=2}^{m-1} A_{k,k+1})A_{m,1} =: A. \tag{63}$$

A solution Z_1 of this equation is called a *generalized m-th root* of the matrix product A.

The m-th roots of matrices are well studied. For a nonsingular matrix A, m-th roots always exist and for a singular matrix A the existence of m-th roots depends on the Jordan structure of A corresponding to the eigenvalue 0, see [31, p. 467].

From Proposition 2.3 we have the following existence result.

Corollary 7.1 *Let $\mathcal{A}$ be as in (60) and let $U = [U_1^H, \ldots, U_m^H]^H$ satisfy $\mathcal{A}U = UR$. If U_1 is nonsingular then the matrices $Z_k = U_{k+1}U_1^{-1}$, $k = 1, \ldots, m-1$, satisfy (61).*

If $\{Z_k\}_{k=1}^{m-1}$ satisfies (61), then the columns of $U = [I, Z_1^H, \ldots, Z_{m-1}^H]^H$ span an invariant subspace of $\mathcal{A}$ corresponding to $R = A_{12}Z_1$.

Clearly if $\{Z_k\}_{k=1}^{m-1}$ satisfy (61) then Z_1 satisfies (63). However, the converse in general does not hold if some $A_{k,k+1}$ is nonsquare or singular. If $m \geq 3$, then the invariant subspace of $\mathcal{A}$ may not lead to all solutions of equation (63). This is the major difference between the problems with $m = 2$ and $m \geq 3$.

Example 7.2 Consider

$$A_{12} = 1, \quad A_{23} = 0, \quad A_{31} = 1.$$

Then $\mathcal{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and equation (63) is scalar, since $z^3 = 0$. So it has only one solution $z = 0$. The equation related to (61) is $z_1^2 = 0$ and $z_2 z_1 = 1$, which clearly has no solution. Note that $\mathcal{A}$ has only one 1-dimensional invariant subspace given by range $[0, 0, 1]^T$, which is just the eigenspace of $\mathcal{A}$.

If all $A_{23}, \ldots, A_{m-1,m}$ are nonsingular, then (61) and (63) are equivalent.

Theorem 7.3 *If $A_{2,3}, \ldots, A_{m-1,m}$ are all nonsingular, then (61) has a solution if and only if (63) has a solution.*

Proof. The necessity is obvious. For the proof of sufficiency let Z_1 be a solution of (63). Then Z_k can be determined recursively via $Z_k = A_{k,k+1}^{-1} Z_{k-1} A_{12} Z_1$, for $k = 2, \ldots, m-1$, and the last equation of (61), $A_{m,1} = Z_{m-1}A_{12}Z_1$, follows from (63). □

If $A_{k,k+1} = I_n$ for $k = 1, \ldots, m-1$ and $A_{m,1} = A \in \mathbf{C}^{n \times n}$, then

$$\mathcal{A} = \begin{bmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & \ddots & I \\ A & & & 0 \end{bmatrix}, \tag{64}$$

and (63) becomes $Z_1^m = A$.

Combining Theorem 7.3 and Corollary 7.1, the matrix m-th root corresponds to an invariant subspace of $\mathcal{A}$. (Note that the condition of Theorem 7.3 is satisfied for this special case.)

Theorem 7.4 *Let $\mathcal{A}$ be as in (64) and let the columns of $U = [U_1^H, \ldots, U_m^H]^H \in \mathbf{C}^{mn \times n}$ span an invariant subspace of $\mathcal{A}$ with $\mathcal{A}U = UR$. If U_1 is nonsingular, then $Z_1 = U_2U_1^{-1} = U_1RU_1^{-1}$ is an m-th root of A, and $Z_1^k = U_{k+1}U_1^{-1}$, for $k = 1, \ldots, m-1$. If Z_1 is an m-th root of A and $U = [I, Z_1^H, \ldots, (Z_1^{m-1})^H]^H$, then $\mathcal{A}U = UZ_1$.*

Note that if Z_1 satisfies (63) then $A_{12}Z_1$ is an m-th root of $A_{12}A$ and Z_1A_{12} is an m-th root of AA_{12}.

Remark 7.5 The matrix sector function can be analyzed in a similar way. Existence and uniqueness of the matrix sector function has been studied in [35].

If

$$\mathcal{A} = \begin{bmatrix} 0 & & & & A \\ A & 0 & & & \\ & A & \ddots & & \\ & & \ddots & & \\ & & & A & 0 \end{bmatrix}, \tag{65}$$

then the m-th sector function S satisfies the invariant subspace relation

$$\mathcal{A} \begin{bmatrix} I \\ S \\ \vdots \\ S^{m-1} \end{bmatrix} = \begin{bmatrix} I \\ S \\ \vdots \\ S^{m-1} \end{bmatrix} AS^{-1}$$

and the polynomial matrix equation

$$(ZA)^{m-1}Z = A^{m-1}. \tag{66}$$

Note that (66) may have many solutions. Solutions exist (even if A is singular), since $Z = I$ is a solution.

8 Numerical methods for general m

For the matrix $\mathcal{A}$ with the block structure in (60) an efficient algorithm can be derived which does not work on the whole matrix $\mathcal{A}$. The following algorithm is a modification of the periodic Schur algorithm of [15, 29, 52].

Algorithm 1.

Input: Matrices $A_{1,2}, \ldots, A_{m-1,m}$, $A_{m,1}$

Output: The Schur form of $\mathcal{A}$ defined in (60).

Let $\mathcal{A} = [A_{i,j}]_{m\times m}$, where $A_{i,j} = 0$ for $i+1 \neq j$ except for $i = m, j = 1$.
Set $\mathcal{U} = I =: [U_{i,j}]_{m\times m}$.

Step 1: Apply the periodic QR algorithm to $A_{1,2}, \ldots, A_{m-1,m}$, $A_{m,1}$, i.e., determine unitary matrices Q_k, $k = 1, \ldots, m$, such that all matrices $A_{k,k+1} := Q_k^H A_{k,k+1} Q_{k+1}$, $k = 1, \ldots, m-1$ and $A_{m,1} := Q_m^H A_{m1} Q_1$ are upper triangular.

Set $\hat{\mathcal{Q}} = \mathrm{diag}(Q_1, \ldots, Q_m)$ and $\mathcal{A} := \hat{\mathcal{Q}}^H \mathcal{A} \hat{\mathcal{Q}}$, $\mathcal{Q} := \mathcal{Q}\hat{\mathcal{Q}}$.

Step 2: **For** $k = 1, \ldots, n$

Let Φ_k be the $m \times m$ matrix

$$\Phi_k = \begin{bmatrix} [A_{11}]_{kk} & \ldots & [A_{1,m}]_{kk} \\ \vdots & \ddots & \vdots \\ [A_{m,1}]_{kk} & \ldots & [A_{m,m}]_{kk} \end{bmatrix}.$$

Determine a unitary matrix P_k such that $P_k^H \Phi_k P_k$ is upper triangular.

Let $\mathcal{P} = [P_{i,j}]$ be the $mn \times mn$ identity matrix except that the k-th diagonal element of block $P_{i,j}$ is replaced by $[P_k]_{i,j}$.

Set $\mathcal{A} := \mathcal{P}^H \mathcal{A} \mathcal{P}$ and $\mathcal{Q} := \mathcal{Q}\mathcal{P}$.

End k

Step 3: **For** $k = 1, \ldots, m-1$

For $\ell = m, \ldots, k+1$,

% Annihilate the block $A_{\ell,k}$

For $i = n-1, \ldots, 1$

For $j = i+1, \ldots, n$

$$\Psi_{i,j} = \begin{bmatrix} [A_{k,k}]_{j,j} & 0 \\ [A_{\ell,k}]_{i,j} & [A_{\ell,\ell}]_{ii} \end{bmatrix}$$

and determine a unitary matrix $W_{i,j}$ such that $W_{i,j}^H \Psi_{i,j} W_{i,j}$ is upper triangular.

Let $\mathcal{W}$ be the identity matrix except for the 2×2 submatrix in the $((k-1)n+j)$-th and $((l-1)n+i)$-th rows and columns which is set to W_{ij}.

Set $\mathcal{A} := \mathcal{W}^H \mathcal{A} \mathcal{W}$ and $\mathcal{Q} := \mathcal{Q}\mathcal{W}$.

End j

End i

End ℓ

End k

Remark 8.1

1. If we apply the Algorithm for the computation of the matrix m-th root, then in Step 1, the periodic Schur decomposition reduces to the classical simple Schur decomposition of A.

2. After Step 2 is completed, all blocks $A_{i,j}$ with $i \leq j$ are upper triangular and all $A_{i,j}$ with $i > j$ are strictly upper triangular.

   The first n columns of $\mathcal{Q}$ span the invariant subspace of $\mathcal{A}$ corresponding to the eigenvalues that appear in the $(1,1)$ entry of $P_k^H \Phi_k P_k$. (For the matrix m-th root, it is convenient here to put the eigenvalue that lies in the first sector, Ω_1, in the $(1,1)$ position of $P_k^H \Phi_k P_k$.)

3. In Step 3, the transformations to eliminate the (i,j) element of $A_{\ell,k}$ does not destroy the triangular form of the blocks. Fill-in is produced only in the (j,i) element of $A_{k,\ell}$.

   If the algorithm is used for computing a matrix m-th root, then one only needs to annihilate $A_{\ell,1}$, $\ell = m, \ldots, 2$ and one only needs to update the first two block rows of $\mathcal{Q}$.

4. A similar algorithm could be used to compute the m-th matrix sector function using the matrix $\mathcal{A}$ (65). This is an unattractive procedure, because the Schur decomposition of A (possibly with some eigenvalue reordering) displays the invariant subspace information of the sector function.

Finally we should point out other matrix equations have similar properties. For example the matrix equation

$$Z^m + A_1 Z^{m-1} + \cdots + A_{m-1} Z + A_m = 0$$

is related to the eigenvalue problem for the block companion matrix

$$\mathcal{A} = \begin{bmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & 0 & I \\ A_m & \ldots & A_2 & A_1 \end{bmatrix}.$$

We are not aware of an efficient method that is able to exploit this structure for computing the Schur form.

9 Conclusion

We have discussed the relation between matrix equations and deflating subspaces of a matrix pencil. The relation covers many important classes of matrix equations including continuous- and discrete-time Riccati equations, Lyapunov, Sylvester and Stein equations as well as matrix m-th roots.

References

[1] F.A. Aliev and V.B. Larin. *Optimization of Linear Control Systems: Analytical Methods and Computational Algorithms.* Stability and Control: Theory, Methods and Applications. Gordon and Breach, 1998.

[2] G.S. Ammar, P. Benner, and V. Mehrmann. A multishift algorithm for the numerical solution of algebraic Riccati equations. *Electr. Trans. Num. Anal.*, 1:33–48, 1993.

[3] J.D. Aplevich. *Implicit Linear Systems.* Lecture Notes in Control and Information Sciences. Springer-Verlag, 1991.

[4] W.F. Arnold, III and A.J. Laub. Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proc. IEEE*, 72:1746–1754, 1984.

[5] Z. Bai and J. Demmel. Design of a parallel nonsymmetric eigenroutine toolbox, Part I. In R.F. Sincovec et al, editor, *Proceedings of the Sixth SIAM Conference on Parallel Processing for Scientific Computing*, 1993. *See also:* Tech. Report CSD-92-718, Computer Science Division, University of California, Berkeley, CA 94720.

[6] R.H. Bartels and G.W. Stewart. Solution of the matrix equation $AX + XB = C$: Algorithm 432. *Comm. ACM*, 15:820–826, 1972.

[7] P. Benner. *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems.* Logos–Verlag, Berlin, Germany, 1997. *Also:* Dissertation, Fakultät für Mathematik, TU Chemnitz–Zwickau, 1997.

[8] P. Benner and R. Byers. Disk functions and their relationship to the matrix sign function. In *Proc. European Control Conf. ECC 97* (CD-ROM), Paper 936, BELWARE Information Technology, Waterloo, Belgium, 1997.

[9] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of embedded Hamiltonian pencils. Technical Report SFB393/99-15, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 1996. Available from `http://www.tu-chemnitz.de/sfb393/sfb99pr.html`.

[10] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical methods for linear-quadratic and H_∞ control problems. In G. Picci and D.S. Gilliam, editors, *Dynamical Systems, Control, Coding, Computer Vision: New Trends, Interfaces, and Interplay*, volume 25 of *Progress in Systems and Control Theory*, pages 203–222. Birkhäuser, Basel, 1999.

[11] P. Benner, V. Mehrmann, and H. Xu. A new method for computing the stable invariant subspace of a real Hamiltonian matrix. *J. Comput. Appl. Math.*, 86:17–43, 1997.

[12] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[13] S. Bittanti, A. Laub, and J. C. Willems, editors. *The Riccati Equation.* Springer-Verlag, Berlin, 1991.

[14] Å. Björck and S. Hammarling. A Schur method for the square root of a matrix. *Linear Algebra Appl.*, 52/53:127–140, 1983.

[15] A. Bojanczyk, G.H. Golub, and P. Van Dooren. The periodic Schur decomposition; algorithms and applications. In *Proc. SPIE Conference, vol. 1770*, pages 31–42, 1992.

[16] S. Boyd, L. El Ghaoui, and E.F.V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* SIAM, Philadelphia, 1999.

[17] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. Numerical methods for algebraic Riccati equations. In S. Bittanti, editor, *Proc. Workshop on the Riccati Equation in Control, Systems, and Signals*, pages 107–116, Como, Italy, 1989.

[18] R. Byers. *Hamiltonian and Symplectic Algorithms for the Algebraic Riccati Equation.* PhD thesis, Cornell University, Dept. Comp. Sci., Ithaca, NY, 1983.

[19] R. Byers. A Hamiltonian QR-algorithm. *SIAM J. Sci. Statist. Comput.*, 7:212–229, 1986.

[20] K.-W.E. Chu. The solution of the matrix equation $AXB - CXD = Y$ and $(YA - CZ, YC - BZ) = (E, F)$. *Linear Algebra Appl.*, 93:93–105, 1987.

[21] F.R. Gantmacher. *Theory of Matrices*, volume 1. Chelsea, New York, 1959.

[22] J.D. Gardiner, A.J. Laub, J.J. Amato, and C.B. Moler. Solution of the Sylvester matrix equation $AXB + CXD = E$. *ACM Trans. Math. Software*, 18:223–231, 1992.

[23] J.D. Gardiner, M.R. Wette, A.J. Laub, J.J. Amato, and C.B. Moler. Algorithm 705: A Fortran-77 software package for solving the Sylvester matrix equation $AXB^T + CXD^T = E$. *ACM Trans. Math. Software*, 18:232–238, 1992.

[24] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Academic Press, New York, 1982.

[25] G. H. Golub, S. Nash, and C. F. Van Loan. A Hessenberg–Schur method for the problem $AX + XB = C$. *IEEE Trans. Automat. Control*, AC-24:909–913, 1979.

[26] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, third edition, 1996.

[27] M. Green and D.J.N Limebeer. *Linear Robust Control.* Prentice-Hall, Englewood Cliffs, NJ, 1995.

[28] B. Hassibi, A.H. Sayed, and T. Kailath. *Indefinite-Quadratic Estimation and Control, a Unified Approach to H^2 and H^∞ Theories.* SIAM, Philadelphia, 1999.

[29] J.J. Hench and A.J. Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39:1197–1210, 1994.

[30] N.J. Higham. Computing real square roots of a real matrix. *Lin. Alg. Appl.*, 88/89:405–430, 1987.

[31] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis.* The Press Syndicate of the University of Cambridge, The Pitt Building, Trumpington Street, Cambridge CB2 1RP, 1991.

[32] V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati theory and robust control. A Popov function approach.* John Wiley & Sons, Chichester, 1999.

[33] C. Kenney and A.J. Laub. The matrix sign function. *IEEE Trans. Automat. Control*, 40(8):1330–1348, 1995.

[34] C. Kenney, A.J. Laub, and M. Wette. Error bounds for Newton refinement of solutions to algebraic Riccati equations. *Math. Control, Signals, Sys.*, 3:211–224, 1990.

[35] C.K. Koç and B. Bakkaloğlu. Halley's method for the matrix sector function. *IEEE Trans. Automat. Contr.*, 40:944–949, 1995.

[36] M.M. Konstantinov, V. Mehrmann, and P.Hr. Petkov. On properties of generalized Sylvester and Lyapunov operators. Preprint 98-12, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 1998.

[37] P. Lancaster. Explicit solutions of linear matrix equation. *SIAM Rev.*, 12:544–566, 1970.

[38] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation.* Oxford University Press, Oxford, 1995.

[39] P. Lancaster and M. Tismenetsky. *The Theory of Matrices.* Academic Press, Orlando, 2nd edition, 1985.

[40] A.J. Laub. A Schur method for solving algebraic Riccati equations. *IEEE Trans. Automat. Control*, AC-24:913–921, 1979.

[41] A.J. Laub. Invariant subspace methods for the numerical solution of Riccati equations. In S. Bittanti, A.J. Laub, and J.C. Willems, editors, *The Riccati Equation*, pages 163–196. Springer-Verlag, Berlin, 1991.

[42] C. Mehl. *Compatible Lie and Jordan algebras and applications to structured matrices and pencils.* Dissertation, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz (FRG), 1998.

[43] C. Mehl. Condensed forms for skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 21:454–476, 1999.

[44] V. Mehrmann. A symplectic orthogonal method for single input or single output discrete time optimal linear quadratic control problems. *SIAM J. Matrix Anal. Appl.*, 9:221–248, 1988.

[45] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[46] V. Mehrmann and H. Xu. Lagrangian invariant subspaces of Hamiltonian matrices. Technical Report SFB393/98-25, Fakultät für Mathematik, TU Chemnitz, 09107 Chemnitz, FRG, 1998. Available from `http://www.tu-chemnitz.de/sfb393/sfb98pr.html`.

[47] T. Penzl. Numerical solution of generalized Lyapunov equations. *Adv. Comp. Math.*, 8:33–48, 1997.

[48] P.Hr. Petkov, N.D. Christov, and M.M. Konstantinov. On the numerical properties of the Schur approach for solving the matrix Riccati equation. *Sys. Control Lett.*, 9:197–201, 1987.

[49] J.D. Roberts. Linear model reduction and solution of the algebraic Riccati equation by use of the sign function. *Internat. J. Control*, 32:677–687, 1980. (Reprint of Technical Report No. TR-13, CUED/B-Control, Cambridge University, Engineering Department, 1971).

[50] G.W. Stewart. *Introduction to Matrix Computations.* Academic Press, New York, 1973.

[51] P. Van Dooren. A generalized eigenvalue approach for solving Riccati equations. *SIAM J. Sci. Statist. Comput.*, 2:121–135, 1981.

[52] P. Van Dooren. Two point boundary value and periodic eigenvalue problems. In O. Gonzalez, editor, *Proc. 1999 IEEE Intl. Symp. CACSD, Kohala Coast-Island of Hawai'i, Hawai'i, USA, August 22–27, 1999* (CD-Rom), pages 58–63, 1999.

[53] H.K. Wimmer. The algebraic Riccati equation without complete controllability. *SIAM J. Alg. Discr. Meth.*, 3:1–12, 1982.

[54] H.K. Wimmer. Decomposition and parametrization of semidefinite solutions of the continuous-time algebraic Riccati equation. *SIAM J. Cont. Optim.*, 32:995–1007, 1994.

[55] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice-Hall, Upper Saddle River, NJ, 1996.